Thermodynamics of adiabatic feedback control.

A.E. Allahverdyan and D.B. Saakian

Yerevan Physics Institute, Alikhanian Brothers Street 2, Yerevan 375036, Armenia.

PACS 05.10.Ln – statistical physics and nonlinear dynamics
PACS 02.30.Yy – control theory
PACS 45.20.Jj – Hamiltonian mechanics

Abstract. - We study adaptive control of classical ergodic Hamiltonian systems, where the controlling parameter varies slowly in time and is influenced by system's state (feedback). An effective adiabatic description is obtained for slow variables of the system. A general limit on the feedback induced negative entropy production is uncovered. It relates the quickest negentropy production to fluctuations of the control Hamiltonian. The method deals efficiently with the entropy-information trade off.

Introduction. Relations between the control theory and physics have a long history. The notorious Maxwell's demon, conceived yet in XIX'th century, is in fact a control device that aims to reduce the entropy of a statistical system [1, 2]. Founders of cybernetics recognized the entropy reduction as one of the basic goals of control [3–5]. This became even more important once it was understood that the statistical description and thermodynamic relations are needed not only for the macroscopic situation, but also for few-body chaotic and stochastic systems [6–9]. Indeed, the first attempts of relating entropy and information during control operations were made in the early days of cybernetics [4, 5] and where based on thermodynamics; see [1] for a fuller historical perspective. More recent theories of entropy-information-control relationship were presented in [2, 10]. The results of [10] found applications in the theory of chaotic systems control, where the entropy reduction is again the basic goal. This field has undergone an explosive development due to synthesing the physical and control scientific ideas [11–13].

Much attention was devoted recently to the control of Brownian systems (mesoscopic particles coupled to a thermal bath) [14]– [25]. This field is expected to have wide applications in various areas of nanoscience. The first theory of feedback driven cooling (entropy-reduction) of a Brownian particle was developed within statistical thermodynamics [14]. This theory plays an important role for recent experimental cooling schemes in nano-physics [15, 16, 23], e.g., in atomic force microscopy [16]. In a related context, Ref. [20] studied experimentally how the feedback control applied to a Brownian nanoparticle generates forces of rather general shape, including non-potential forces.

Following to the experimental development of feedback cooling methods, Ref. [17] presented the thermodynamic analysis of a Brownian particle, which couples to a thermal bath and is manipulated by control fields so that to cool the bath. The authors of [17] gave a general recipe for calculating the entropy pumped out of the bath versus the entropy produced during the operation of the Brownian particle. The quantum extension of this setup was investigated in [18]. Fluctuation theorems for the classical Brownian control setup were studied in [19].

Control of Brownian particles is also employed for generating a directed motion; see [21] for reviews. This task is important for contructing nanoscale engines (rachets or Brownian motors). Theoretical and experimental proposals for feedback driven ratchets appreared recently in [22] and [23], respectively.

Finally we should mention works devoted to the open-loop (non-feedback) control of Brownian particles [24, 25]. These studies are especially relevant for statistical physics, since the basic formulations of the second law are in fact control-theoretical statements [26].

Here we explore an approach to the adaptive feedback control of classical Hamiltonian systems. Feedback means that the dynamics of a control parameter of the Hamiltonian is influenced by system's state, i.e., it performs an adaptive motion, while in the non-feedback (open-loop) control the motion of the control parameter is prescribed. Our main assumption is that the control parameter moves slowly. Assuming the ergodicity of certain observables we develop a general thermodynamic description of the feed-

back control process. In particular, we find the control fields that ensure the quickest reduction of the entropy (noise). This reduction is related to fluctuations of the controlling part of the Hamiltonian. We also describe the entropy-information trade-off: how limitations of the information available on system's state decrease the speed of the entropy reduction and change the qualitative features of the control process.

Note that in contrast to the above works on the Brownian particles control, we shall work with the full Hamiltonian system, and not with a small particle coupled to a large thermal bath. Moreover, we focus on Hamiltonian systems with finite degrees of freedom, though the presented theory applies to macrocopic Hamiltonian systems, e.g., the particle plus the bath. The macroscopic system control will be studied elsewhere [34].

Basics of the method. Consider a system with n degrees of freedom and Hamiltonian $H(p,q,R)$. The equations of motion read

$$\dot{q} = \partial_p H(p,q,R), \qquad \dot{p} = -\partial_q H(p,q,R), \tag{1}$$

where $q=(q_1,...,q_n)$ and $p=(p_1,...,p_n)$ are, respectively, the coordinates and momenta, and where R is a time-varying parameter (the extension to several parameters is straightforward). The parameter R is changed externally controlling the system (the goals of control are indicated below). The evolution of R is described in the standard way of the adaptive control [11–13, 27, 28]:

$$\tau_R \dot{R} = F(E,R,z), \quad z \equiv (q,p), \quad |F| \leq \widetilde{F}, \tag{2}$$

where F is the control field, τ_R is a characteristic time of R, and where E is system's energy. The constraint $|F| \leq \widetilde{F}$ with constant $\widetilde{F}$ is a natural condition for practical realizability of the control setup. Many control tasks get the real physical meaning only after imposing such constraints on the magnitude of the control fields [27].

As shown by (2), the controlling parameter R is subjected to feedback: the dynamical variables z and E influence, via suitable engineering, the evolution of R. Control processes without feedback (open-loop control) correspond to $F=F(R)$, with predetermined motion of R.

So far we presented a standard and rather general setup for control problems. In particular, Eq. (2) contains many adaptive control processes known in literature [11–13, 27, 28]. One of our basic assumptions is that the motion of R is adiabatic, i.e., that the time-scale τ_R of R is much larger than the characteristic time τ_S of the system (defined after (10)). Eqs. (1, 2) show that together with R also the energy E becomes a slow variable,

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\mathrm{d}R}{\mathrm{d}t}\,\partial_R H(z,R) = \frac{1}{\tau_R} F(E,R,z)\partial_R H(z,R), \tag{3}$$

provided that the controlling part of the Hamiltonian $\partial_R H$ is limited. Note that $\frac{\mathrm{d}R}{\mathrm{d}t}$ is equal to the small parameter $\frac{1}{\tau_R}$ times the function $F(E,R,z)$, which changes fast together with z. Thus we need the constraint (2) on the magntitude of $F(E,R,z)$ for the adiabatic assumption to apply.

We want to get from (1–3) averaged equations for the slow variables E and R. To this end, define the microcanonic distribution:

$$\mathcal{M}(z,E,R) = \frac{\delta[E - H(z,R)]}{\partial_E \Omega(E,R)}, \tag{4}$$

$$\Omega(E,R) = \int \mathrm{d}z\, \theta[E - H(z,R)], \tag{5}$$

and where $\delta(x)$ and $\theta(x)$ are, respectively, the delta and step function. Here $\Omega(E,R)$ is the phase-space volume at energy E; its derivative $\partial_E \Omega(E,R)$ defines the normalization of $\mathcal{M}(z,E,R)$.

Denote by z_t and R_t the solutions of (1, 2). On times $\tau_R \gg \tau \gg \tau_S$ we have from (1, 2) for the slow derivative $\mathbf{d}E/\mathbf{d}\tau$:

$$\begin{aligned}
\frac{\mathbf{d}E}{\mathbf{d}\tau} &\equiv \frac{1}{\tau}\left[\, H(z_{t+\tau}, R_{t+\tau}) - H(z_t, R_t)\,\right] \qquad (6)\\
&= \int_t^{t+\tau} \frac{\mathrm{d}s}{\tau}\frac{\mathrm{d}H}{\mathrm{d}s}(z_s,R_s) = \int_t^{t+\tau}\frac{\mathrm{d}s}{\tau}\,\dot{R}_s\,\frac{\partial H}{\partial R}(z_s,R_s),\\
&= \frac{1}{\tau_R}\int_t^{t+\tau}\frac{\mathrm{d}s}{\tau}\,F(E_s,R_s,z_s)\frac{\partial H}{\partial R}(z_s,R_s),\\
&= \frac{1}{\tau_R}\int_t^{t+\tau}\frac{\mathrm{d}s}{\tau}\,F(E_t,R_t,z_s)\frac{\partial H}{\partial R}(z_s,R_t) + o(\frac{\tau}{\tau_R}). \qquad (7)
\end{aligned}$$

As a consequence of the adiabatic assumption, the last integral in (7) refers to the dynamics with R_t =const and E_t =const. Denote

$$w(z) \equiv F(E_t,R_t,z)\,\partial_R H(z,R_t), \tag{8}$$

and write the time-average in (7) as $\int_t^{t+\tau}\frac{\mathrm{d}s}{\tau}\,w(z_s)$. Consider the following obvious relation:

$$\int \mathrm{d}z\, w(z)\mathcal{M}(z,E) = \frac{1}{\tau}\int_t^{t+\tau}\mathrm{d}s\int \mathrm{d}z\, w(z)\mathcal{M}(z,E). \tag{9}$$

In the RHS of (9) we change the integration variable as $y = \mathcal{T}_{t-s}\,z$, where $\mathcal{T}_t$ is the flow generated by the Hamiltonian $H(z)$ between times 0 and t. Employing Liouville's theorem, $\mathrm{d}z = \mathrm{d}y$, and energy conservation, $\mathcal{M}(z,E) = \mathcal{M}(y,E)$, one gets

$$(9) = \int \mathrm{d}y\,\mathcal{M}(y,E)\,\frac{1}{\tau}\int_t^{t+\tau}\mathrm{d}s\, w(\mathcal{T}_{s-t}\,y). \tag{10}$$

If $w(z)$ is an *ergodic observable* of the R_t =const dynamics, then for $\tau \gg \tau_S$ the time-average $\frac{1}{\tau}\int_t^{t+\tau}\mathrm{d}s\,w(\mathcal{T}_{s-t}\,y)$ in (10) depends on the initial condition y only via its energy $H(y)$ [6, 7, 32]. (Thus, τ_S is the relaxation time of $w(z)$.) In particular, the dependence of the precise value of y is irrelevant provided that the condition $H(y)=E$ holds. Since $\mathcal{M}(y,E)$ is a delta-function concentrated at that value of energy, the integration over y in (10) drops

out, and we get from (9, 10) that the time-average is equal to the microcanonic average at the energy E

$$\int \mathrm{d}z\, w(z)\mathcal{M}(z,E) = \frac{1}{\tau}\int_t^{t+\tau} \mathrm{d}s\, w(\mathcal{T}_{s-t}\, y). \tag{11}$$

Applying this to (7) we get

$$\tau_R \frac{\mathbf{d}E}{\mathbf{d}\tau} = \int \mathrm{d}z\ \mathcal{M}(z,E,R)\, F(E,R,z)\, \partial_R H(z,R), \tag{12}$$

$$\tau_R \frac{\mathbf{d}R}{\mathbf{d}\tau} = \int \mathrm{d}z\ \mathcal{M}(z,E,R)\, F(E,R,z) \equiv \langle F\rangle_{E,R}, \tag{13}$$

where (13) is derived analogously to (12) by assuming the ergodicity of $F(E,R,z)$.

Eqs. (12, 13) are the basic equations of the adiabatic feedback theory. We list again the assumptions employed in their derivation: *i)* E and R are slow variables; *ii)* conservation of energy for $R=\mathrm{const}$; *iii)* Liuoville's theorem; *iv)* ergodicity of two phase-space observables: $w(z)$ defined by (8) and the controlling parameter $F(E,R,z)$.

Instead of $2n+1$ equations (1, 2) we have in (12, 13) only two equations for E and R. They are autonomous, since they do not depend on the precise initial value of z, provided it was on the initial energy shell E_i. Thus the control processes described by (12, 13) can operate under conditions, where the initial values of z are not known or the dependence from them is not desirable. The price to be paid for this is that now only functions of E and R can be controlled.

Recall that for a (fully) ergodic system all the sufficiently smooth observables are ergodic, while a non-ergodic system can still have some ergodic observables; see [32] for the general theory. Now note that since no ergodicity of all observables is assumed in deriving Eqs. (12, 13), they apply to some non-ergodic systems. Another reason for applying (12, 13) to non-ergodic systems is that the ergodicity may be restored under small perturbation [6,7]. Thus the scheme applies to most of chaotic systems.

Definition of entropy. Entropy and information play important roles in the control theory: the very possibility of applying feedback is due to the information available on the state of the system, while entropy reduction [negentropy production] is necessary for immunizing the system from sources of noise and instability [2–4]. Thus our next task is to obtain from (12, 13) the maximal negentropy production rate. Recall that for a (partially) ergodic system the entropy is defined as [6,7]:

$$S = \ln\Omega(E,R) \equiv \ln\int \mathrm{d}z\, \theta(E-H(z,R)). \tag{14}$$

Eq. (14) satisfies to all desired features of entropy, even if the number n of the system degree of freedom is finite:

1. For the temperature T defined via the standard thermodynamic formula

$$1/T = \beta = \partial_E \ln\Omega(E) = \frac{1}{\Omega(E)}\int \mathrm{d}z\, \delta(E-H(z,R)), \tag{15}$$

the integration by parts leads to equipartition [6, 7]: $\langle y\, \partial_y H\rangle = T$, where y is any canonical variable, while $\langle ...\rangle$ is the average over microcanonic distribution (5). For the standard Hamiltonian $H=\sum_{k=1}^n \frac{p_k^2}{2} + V(q_1,..,q_n)$ we get $\langle p_k^2\rangle = T$ for any k, which is the standard form of equipartition. Note that T in (15) is non-negative.

2. S satisfies to the first law of thermodynamics for the microcanonic ensemble [6,7,30].

3. S satisfies to two formulations of the second law that describe a partially ergodic system subjected to an open-loop (i.e., non-feedback) control: *i)* S remains invariant in an adiabatically slow [open-loop] process; see [6–9] and the discussion after (17). *ii)* S increases under a non-slow [open-loop] process, provided that the system starts its evolution from an equilibrium state [8, 31]. The latter formulation is closely related to the minimum work principle [26].

In the thermodynamical limit S goes to the more usual Boltzmann expression [6,7]:

$$S_B(E) = \ln[\partial_E\Omega] = \int \mathrm{d}z\, \delta(E-H(z,R)). \tag{16}$$

None of the above features **1-3** holds if we apply S_B out of the thermodynamic limit, e.g., S_B is not constant for open-loop adiabatic processes. This is because S is the unique adiabatic invariant for ergodic systems; see [6, 7] and the discussion after (17). Another problem with using the Boltzmann expression S_B directly for finite systems is that the temperature defined via (16) and the standard thermodynamic formula as $1/T_B = \partial_E S_B(E)$ is in general not even positive [7]. The problems in attempting to use S_B as the proper entropy will be illustrated below by a concrete example; see the discussion after (34).

We are thus convinced that S is the proper expression of entropy for both finite and macroscopic ergodic systems.

Negentropy production. We now determine the evolution of the entropy S according to (12, 13). Using (12–15) we get

$$\tau_R \frac{\mathbf{d}S}{\mathbf{d}\tau} = \beta(E,R)\, \langle\, F\, [\, \partial_R H - \langle \partial_R H\rangle_{E,R}\,]\, \rangle_{E,R}, \tag{17}$$

where $\langle ...\rangle_{E,R}$ is defined in (13), and where $\beta(E,R) > 0$ is the inverse temperature defined in (15). Eq. (17) shows that if there is no feedback over the fast variable z, $F=F(E,R)$, the entropy $S=\ln\Omega(E,R)$ is conserved on the slow time, i.e., it is an adiabatic invariant [6–9]. Recall that for ergodic systems this is the unique adiabatic invariant: any other quantity $K=K(E,R)$ that remains constant for open-loop adiabatic processes is a function of the phase-space volume Ω: $K=K(\Omega)$ [6,7]. Indeed, since in general $\partial_E\Omega(E,R) > 0$ we can express (for a fixed R) E as a function of Ω and R: $E=E(\Omega,R)$. Putting this into $K(E,R)$ we re-express it as a function of Ω and R: $K=K(\Omega,R)$. Differentiating $K(\Omega,R)$ over the slow time we get:

$$\frac{\mathbf{d}K}{\mathbf{d}\tau} = \frac{\mathbf{d}\Omega}{\mathbf{d}\tau}\,\frac{\partial K}{\partial\Omega} + \frac{\mathbf{d}R}{\mathbf{d}\tau}\,\frac{\partial K}{\partial R}. \tag{18}$$

Since both Ω and K are assumed to be adiabatic invariants, $\frac{\mathbf{d}\Omega}{\mathbf{d}\tau} = \frac{\mathbf{d}K}{\mathbf{d}\tau} = 0$, we get from (18) that $\frac{\partial K}{\partial R} = 0$, i.e., K is a function of Ω only.

However, for feedback processes the entropy does change. Let us find the feedback $F(E,R,z)$ that maximizes the negentropy production $-\frac{\mathbf{d}S}{\mathbf{d}\tau}$ under the natural constraint $|F| \leq \widetilde{F}$. Since the RHS of (17) is a linear function of F, and since $\beta > 0$, the extremum is achieved on the boundaries $|F| = \widetilde{F}$ of the allowed range. This implies for the most negative negentropy production $\frac{\mathbf{d}S}{\mathbf{d}\tau}$, which we denote as $\widetilde{\frac{\mathbf{d}S}{\mathbf{d}\tau}}$:

$$\tau_R \widetilde{\frac{\mathbf{d}S}{\mathbf{d}\tau}} = -\beta\, \widetilde{F} \langle\, |\partial_R H - \langle \partial_R H \rangle_{E,R}|\, \rangle_{E,R}, \quad (19)$$

$$F(E,R,z) = -\widetilde{F}\, \mathrm{sign}[\, \partial_R H - \langle \partial_R H \rangle_{E,R}\,]. \quad (20)$$

Eq. (19) bounds the rate of the entropy reduction, and it is related to fluctuations of the control Hamiltonian $\partial_R H$ on the surface of constant energy. The optimal control function (20) is seen to be discontinuous. Note that Ref. [25] reports that discontinuous control fields is a general feature of the optimal open-loop control that operates in a finite time. We see a similar effect for a feedback setup, which is not constrained globally to a finite operation time.

The phase-space volume $\Omega(E,R)$ is a Lyapunov function for the dynamics described by (12, 13, 20), since it is obviously non-negative, and since it is non-increasing, $\tau_R \widetilde{\frac{\mathbf{d}\Omega}{\mathbf{d}\tau}} \leq 0$, as follows from (19). The non-negativity and non-increasing of $\widetilde{\frac{\mathbf{d}\Omega}{\mathbf{d}\tau}}$ imply, via the Lasalle principle [29], that the long-τ solutions $\bar{E}$, $\bar{R}$ of (12, 13, 20) satisfy $\widetilde{\frac{\mathbf{d}\Omega}{\mathbf{d}\tau}}(\bar{E},\bar{R}) = 0$, which leads via (19) to

$$\int \mathrm{d}z\, \delta(\bar{E} - H(z,\bar{R}))[\partial_R H(z,\bar{R}) - \langle \partial_R H \rangle_{\bar{E},\bar{R}}] = 0. \quad (21)$$

There are two possibilities for satisfying the equality in (21): *i)* the long-τ solutions converge to a stable fixed point (i.e., energy minimum) of the original Hamiltonian system (1). This means that the phase-space volume $\Omega(E,R)$ decays to zero, while the entropy $\ln\Omega(E,R)$ decays to its minimal value minus infinity. *ii)* The second possibility of realizing (21) is that the long-τ solutions converge to a point $(\bar{E},\bar{R})$ such that $\partial_R H(z,\bar{R})$ as a function of z is constant on the energy shell $H(z,\bar{R}) = \bar{E}$, i.e., it is a constant of motion for the fixed values of $E = \bar{E}$ and $R = \bar{R}$. Since the second option is unstable to small perturbation in the control Hamiltonian $\partial_R H(z,R)$, the first option is more likely to be realized: the feedback setup (20) drives the system toward an energy minimum, where the phase-space volume Ω is equal to zero.

Limits on the available information. When obtaining the maximum rate (19) of negentropy production we only assumed that the magnitude $|F|$ of the controlling parameter is limited from above; see (2). More crucial restrictions come into play when noting that in practice the very information available on the state of the system is limited. We thus assume that the full knowledge of z is not available for the controller; only some function $\Phi(z)$ of z is known, and thus the feedback F in (2) depends on z only via $\Phi(z)$:

$$F(E,R,z) = f(E,R,\Phi(z)). \quad (22)$$

Note that this implies a reduction of the data z, since for simplicity we take one—in general not one-to-one—function $\Phi(z)$ instead of the vector $z = (q_1,...,q_n; p_1,...,p_n)$. This is the standard way of modeling the data reduction in the information theory, known also as coarse-graining or statistics taking [33]. All the standard measures of information—e.g., Shannon entropy, relative entropy, the Fisher information—are known to decrease after taking a (not one-to-one) function of the data. In other words, the data reduction means information decrease with respect to any definition of information. In the extreme case, where no information is available for the feedback we have $\Phi(z) = \mathrm{const}$. Rewriting (17) as

$$\tau_R \frac{\mathbf{d}S}{\mathbf{d}\tau} = \beta(E,R) \int \mathrm{d}y\, f(E,R,y)\, \psi(E,R,y), \quad (23)$$

where we defined

$$\psi(E,R,y) \equiv \int \mathrm{d}z\, \mathcal{M}(z,E,R) \times \quad (24)$$
$$\delta[y - \Phi(z)]\, [\, \partial_R H(z,R) - \langle \partial_R H \rangle_{E,R}\,],$$

and applying the same derivation as for (20, 19), we get for the most negative $\frac{\mathbf{d}S}{\mathbf{d}\tau}$ at the given $\Phi(z)$:

$$\frac{\mathbf{d}S}{\mathbf{d}\tau} = -\frac{\beta \widetilde{F}}{\tau_R} \int \mathrm{d}y\, |\, \psi(E,R,y)\, |\, . \quad (25)$$

This value of $\frac{\mathbf{d}S}{\mathbf{d}\tau}$ is achieved for the feedback:

$$f(E,R,y) = -\widetilde{F}\, \mathrm{sign}[\, \psi(E,R,y)\,], \quad (26)$$

where $F(E,R,z)$ is recovered from (22, 26). Thus the maximal negentropy production under information limitation is related to fluctuations of the control Hamiltonian $\partial_R H$ over a constrained microcanonic ensemble; see (24, 25). As follows from (17, 25), the speed of entropy reduction decreases under information limitations:

$$|\frac{\mathbf{d}S}{\mathbf{d}\tau}| \leq |\widetilde{\frac{\mathbf{d}S}{\mathbf{d}\tau}}|. \quad (27)$$

In particular, $\psi = \frac{\mathbf{d}S}{\mathbf{d}\tau} = 0$ for $\Phi = \mathrm{const}$.

Examples. We illustrate the obtained feedback schemes via the celebrated example of adiabatic physics that is a harmonic oscillator with the controlling frequency:

$$H = \frac{p^2}{2} + \frac{Rq^2}{2}. \quad (28)$$

This Hamiltonian with the feedback controlling frequency is close to the experimental situation realized in Ref. [20]. In the context of adiabatic feedback control, the harmonic

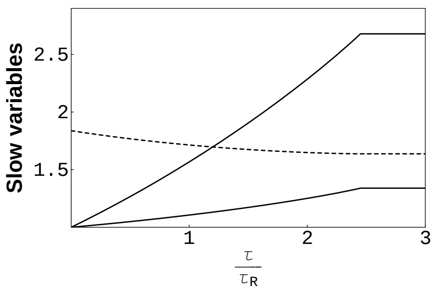

Fig. 1: Entropy reduction scheme (33, 34) for harmonic oscillator. Slow variables versus dimensionless time τ/τ_R: Energy E (lower curve), frequency R (upper curve), and entropy S (dashed curve). For the parameters we take $\widetilde{F}=1$ and $L=1$. The curves saturate for $\tau/\tau_R > 2.5$ when the available information is not anymore sufficient for any entropy reduction.

oscillator (28) displays two interesting effects that exist as well in more elaborated situations [34]: control without systematic motion of the controlling parameter and qualitative changes in the control setup upon information limitations. For a constant R the period of the oscillator is $2\pi/\sqrt{R}$, and the adiabatic approach applies at least for $1 \ll \tau_R\sqrt{R}/2\pi$. This is an ergodic system and Eq. (20) implies for the optimal entropy-reducing control

$$F(E,R,q) = -\frac{\widetilde{F}}{2}\,\mathrm{sign}(q^2 - \frac{E}{R}). \tag{29}$$

Eq. (29) leads via (12, 13, 5) to R constant in the slow time (though R is not constant on the fast time),

$$\frac{\mathbf{d}R}{\mathbf{d}\tau} = 0, \tag{30}$$

exponential decay of energy (cooling),

$$E(\tau) = E(0)e^{-\frac{\widetilde{F}}{\pi R}\frac{\tau}{\tau_R}}, \tag{31}$$

and thus to linear decay of the entropy $S = \ln[\frac{2\pi E}{\sqrt{R}}]$: $S(\tau) = S(0) - \frac{\widetilde{F}}{\pi R}\frac{\tau}{\tau_R}$. As intuitively expected, entropy reduction relates to cooling. Eq. (30) shows that the controlling parameter R does not move in average, i.e., on the slow time. The cooling is achieved due to the motion of R on the fast time-scale; see (2, 29).

To limit the information available about the coordinate q, we assume that it is only known whether q is larger or smaller than a positive constant L. The function $\Phi(q)$ in (22) thus takes only two distinct values, e.g., $\Phi(q) = \theta(L-q)$. This brings from (26, 22) a control setup

$$F(E,R,q) = \widetilde{F}\,\mathrm{sign}(L-q)\,\theta(2E - RL^2). \tag{32}$$

The latter step function is there, since for small E (or large R) the oscillator is located next to $q=0$ and its position cannot be utilized by the feedback. Thus for those values of E it is best to do nothing, $F=0$, since any control (under assumed information limits) will increase the entropy (disorder). We get from (12, 13, 32) and from (22–25)

$$\frac{\mathbf{d}R}{\mathbf{d}\tau} = \frac{2\widetilde{F}}{\tau_R\pi}\,\theta(1-\xi)\,\arcsin\xi, \qquad \xi \equiv \sqrt{\frac{RL^2}{2E}}, \tag{33}$$

$$\frac{\mathbf{d}S}{\mathbf{d}\tau} \equiv \frac{\mathbf{d}}{\mathbf{d}\tau}\ln[\frac{2\pi E}{\sqrt{R}}] = -\frac{\widetilde{F}}{\pi R\tau_R}\,\theta(1-\xi)\,\xi\sqrt{1-\xi^2}, \tag{34}$$

while the equation for $\frac{\mathbf{d}E}{\mathbf{d}\tau}$ can be recovered from (33, 34). The behavior of $E(\tau)$, $R(\tau)$ and $S(\tau)$ is shown in Fig. 1. We see that the entropy reduction rate is not simply smaller than the optimal one, but it is realized via energy increase (heating) rather than cooling.

For the considered oscillator model, let us illustrate that the Boltzmann expression is not the proper defintion of entropy for a finite system. Recalling (14, 16) and using (34) we get that for the harmonic oscillator (28): $S_B = \ln[\frac{2\pi}{\sqrt{R}}]$. It is seen that *i)* S_B does not depend on the energy, so that attempting to define the temperature via the standard formula (15) will lead to zero temperature, not a reasonable result. *ii)* S_B is not adiabatic invariant, one can decrease it via an open-loop control by changing R.

Adiabatic invariant. Eq. (32) provides a control setup, where the dependence on the slow and fast variables factorizes:

$$F(E,R,z) = g(E,R)\phi(z). \tag{35}$$

For this case the slow variable system (12, 13) possesses an integral of motion, i.e., an adiabatic invariant. One deduces from (12, 13):

$$\frac{\mathbf{d}}{\mathbf{d}\tau}\int \mathrm{d}z\,\phi(z)\,\theta(E - H(z,R)) = 0. \tag{36}$$

This conservation generalizes to many-dimensional systems the adiabatic invariance found in [35]. For $\phi(z) =$ const, where there is no feedback over fast variables, we are naturally back to the usual conservation of the phase-space volume.

In conclusion, we developed an adiabatic approach for the adaptive feedback control of Hamiltonian systems. It is derived assuming ergodicity of two observables and thus applies to the most of chaotic systems and some integrable ones. The approach reduces the control problem to two equations (12, 13) describing the evolution of slow variables. With help of these equations we got a general upper bound (19) on the rate of negentropy (order) production induced by the feedback control. This bound is achieved for discontinuous control field (20), and is related to the fluctuation of the control Hamiltonian over the microcanonic ensemble.

The method describes the information-entropy trade-off: how the maximal negentropy production rate decreases when the information available for the feedback gets limited. The example of harmonic oscillator with the

controlling frequency shows that information limitations do change the quanlitative features of the control dynamics. In particular, the entropy reduction is realized via heating the system. Note that in the present approach we standardly modeled the information limitation via the reduction of the data available to the controller.

The Hamiltonian dynamics finds applications well beyond the proper (statistical) mechanics, e.g., in hydrodynamics [7] or in ecology [36]. Control issues in these fields are well known, and since our methods are not system-specific, they may apply to controlling a vortex flow, or to optimizing the harvest production [34].

* * *

We thank K.G. Petrosyan for useful discussions.

The work was supported by Volkswagenstiftung (grant "Quantum thermodynamics: energy and information flow at nanoscale") and by CRDF Grant No. ARP2-2647-YE-05.

REFERENCES

[1] H.S. Leff and A.F. Rex, *Maxwell's Demon: Entropy, Information, Computing* (IOP, Bristol, 2003).
[2] R. Poplavskii, Sov. Phys. Usp. **22**, 371 (1979).
[3] N. Wiener, *Cybernetics* (MIT Press, Cambridge, 1948).
[4] W.R. Ashby, *An Introduction to Cybernetics* (Wiley, NY, 1956).
[5] L. Brillouin, *Science and Information Theory* (Academic Press, NY, 1956).
[6] A. Munster, *Statistical Thermodynamics* (Springer-Verlag, Berlin, 1969), Vol. 1.
[7] V. L. Berdichevsky, *Thermodynamics of Chaos and Order* (Addison Wesley Longman, Essex, England, 1997).
[8] S. Sasa and T.S. Komatsu, Prog. Theor. Phys. **103**, 1 (2000).
[9] H.H. Rugh, Phys. Rev. E **64**, 055101 (2001).
[10] H. Touchette and S. Lloyd, Phys. Rev. Lett. **84**, 1156 (2000); Physica A **331**, 140 (2004).
[11] D.J. Gauthier, Am. J. Phys. **71**, 750 (2003). S. Boccaletti *et al*, Phys. Rep. **329**, 103 (2000).
[12] J. Bechhoefer, Rev. Mod. Phys. **77**, 783 (2005).
[13] A.L. Fradkov, Physics Uspekhi, **48**, 113 (2005).
[14] C. Kittel, *Elementary Statistical Physics* (Wiley, NY, 1958).
[15] B. D'Urso *et al.*, Phys. Rev. Lett. **90**, 043001 (2003). P. F. Cohadon *et al.*, Phys. Rev. Lett. **83**, 3174 (1999). P. Bushev *et al.*, Phys. Rev. Lett. **96**, 043003 (2006).
[16] S. Liang *et al.*, Ultramicroscopy **84**, 119 (2000).
[17] K.H. Kim and H. Qian, Phys. Rev. Lett. **93**, 120602 (2004).
[18] M. Pavon and F. Ticozzi, J. Math. Phys. **47**, 063301 (2006).
[19] K.H. Kim and H. Qian, Phys. Rev. E **75**, 022102 (2007).
[20] A.E. Cohen, Phys. Rev. Lett. **94**, 118102 (2005); PNAS **103**, 4362 (2006).
[21] J.M.R. Parrondo and B. Jimenez de Cisneros, Appl. Phys. A **75**, 179 (2002). P. Reimann, Phys. Rep. **361**, 57 (2002).
[22] F.J. Cao, L. Dinis and J.M.R. Parrondo, Phys. Rev. Lett., **93**, 040603 (2004). L. Dinis, J.M.R. Parrondo1 and F.J. Cao, Europhys. Lett., **71**, 536 (2005).
[23] V. Serreli et al., Nature **445**, 523 (2007).
[24] K. Sekimoto and S. Sasa, J. Phys. Soc. Jpn. **66**, 3326 (1997).
[25] T. Schmiedl and U. Seifert, Phys. Rev. Lett. **98**, 108301 (2007).
[26] A.E. Allahverdyan and Th.M. Nieuwenhuizen, Phys. Rev. E, **75**, 051124 (2007); *ibid.*, **71**, 046107 (2005).
[27] A.G. Butkovskiy and Yu.I. Samoilenko, *Control of Quantum-Mechanical Processes and Systems* (Kluver, Dordrecht, 1990).
[28] B.A. Huberman and E. Lumer, IEEE Trans. Circuits Syst., **37**, 547 (1990); S. Sihna *et al*, Physica D **43**, 118 (1990); J.K. John and R.E. Amritkar, Phys. Rev. E **49**, 4843 (1994); D. Huang, Phys. Rev. E. **73**, 066204 (2006).
[29] J.P. Lasalle, IRE Trans. Circuit Theory **7**, 520 (1960).
[30] M. Campisi, Stud. Hist. Phil. M. P. **36**, 275 (2005).
[31] M.Campisi, cond-mat/07042567.
[32] N. G. van Kampen, Physica **53**, 98 (1971).
[33] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley, New York, 1991).
[34] A.E. Allahverdyan, in preparation.
[35] P. Helander *et al.*, Phys. Rev. Lett. **68**, 3659 (1992).
[36] E.H. Kerner, J. Math. Phys. **38**, (1997).